Survival Models of Community Tenure and Length of Hospital Stay for the Seriously Mentally Ill: A 10-year Perspective

Steven Stern
University of Virginia

Elizabeth Merwin
University of Virginia

Fredrick Holt
Bates, White, and Ballentine, LLC

November, 2000

Revised May, 2001

Revised September, 2001

The corresponding author is Steven Stern, Department of Economics, University of Virginia, Charlottesville, VA 22903. The data and tabulations utilized in this paper were available through the Virginia DMHMRSAS. The data were originally collected by state mental health facilities. Neither the original source nor collectors of the data maintain any responsibility for the analyses or interpretations presented here. This project was supported by NIMH R29 MH48968, NIMH R01 MH53259-01A2, and the Southeastern Rural Mental Health Research Center grant NIMH P50MH49173.

Abstract

Objective: To examine the effects of personal and community characteristics, specifically race and rurality, on lengths of state psychiatric hospital and community stays using maximum likelihood survival analysis with a special emphasis on change over a ten year period of time.

Data Sources: We used the administrative data of the Virginia Department of Mental Health, Mental Retardation, and Substance Abuse Services (DMHMRSAS) from 1982-1991 and the Area Resources File (ARF). Given these two sources, we constructed a history file for each individual who entered the state psychiatric system over the ten year period. Histories included demographic, treatment, and community characteristics.

Study Design: We used a longitudinal, population-based design with maximum likelihood estimation of survival models. We presented a random effects model with unobserved heterogeneity that was independent of observed covariates. The key dependent variables were lengths of inpatient stay and subsequent length of community stay. Explanatory variables measured personal, diagnostic, and community characteristics, as well as controls for calendar time.

Data Collection: This study used secondary, administrative and health planning data.

Principal Findings: African-American clients leave the community more quickly than whites. After controlling for other characteristics, however, race does not affect hospital length of stay. Rurality does not affect length of community stays once other personal and community characteristics are controlled for. However, people from rural areas have longer hospital stays even after controlling for personal and community characteristics. The effects of time are significantly smaller than expected. Diagnostic composition effects and a decrease in the rate of first inpatient admissions explain part of this reduced impact of time. We also find strong evidence for the existence of unobserved heterogeneity in both types of stays and adjust for this in our final models.

Conclusions: Our results show that information on client characteristics available from inpatient stay records is useful in predicting not only the length of inpatient stay but also the length of the subsequent community stay. This information can be used to target increased discharge planning for those at risk of more rapid readmission to inpatient care. Correlation across observed and unobserved factors affecting length of stay has significant effects on the measurement of relationships between individual factors and lengths of stay. Thus, it is important to control for both observed and unobserved factors in estimation.

Keywords: community tenure, length of psychiatric inpatient stay, survival analysis, state psychiatric hospital, maximum likelihood estimation

1. Introduction

The successful transition from state owned and operated inpatient mental health care to community care has been an ongoing challenge for many states. Indeed, reducing admissions, reducing lengths of stay, and reducing re-admissions have been a constant clinical and administrative challenge since the 1960s. Despite hundreds of studies of these topics, these issues continue as barriers to successful community treatment and community living. Early research in this area had limited usefulness due to a lack of data, the complexity of what data there were, and a lack of appropriate methods for studying the issues. This study uses approximately 10 years of data, including information on care given to all clients admitted during that time period to any state mental hospital in

Virginia. This time span allows the influence of time on community tenure to be fully explored. Our study presents a random effects model with unobserved heterogeneity independent of observed covariates.

We consider the influence of patient characteristics and community characteristics on the probabilities of leaving the hospital and of exiting the community. (i.e. returning to the hospital). Using econometric methods, we account for the influence of non-measured variables such as severity of illness on the outcomes of interest, and we provide important explanations of community tenure over a 10-year period. The methods employed capitalize on the sophistication of econometrics and take advantage of the more recent availability of many longitudinal administrative data sets. We offer the field of health care study a new means for analyzing on old issue. Despite extensive study of community tenure, factors predictive of successful community living following inpatient care are still only partially understood. Particularly lacking is an understanding of the use of inpatient care in facilitating community care for those with serious mental illness.

We develop several models explaining community tenure and the duration of inpatient care stays and present survival models for different types of vulnerable subgroups of patients. To facilitate replication, we describe our methods in some detail.

2. Relevant Literature

From 1980 to 1994, the number of state and county mental hospitals dropped from 280 to 256 (Witkin 1998, p. 145). Their bed numbers dropped from 156,482 to 81,911 during that time period. Thus, in the aggregate, the health policy of deinstitutionalization of the seriously mentally ill has been successful in reducing inpatient care provided by states and counties. State owned and operated psychiatric beds per 100,000 people declined from 70 to 32 during this time period (p. 146). Nationally, there was a corresponding increase in the number of organizations providing out-patient and/or partial care (less-than-24-hour care) from 2,431 in 1980 to 4,087. The goal has been to shift inpatient care to the community (p. 145).

Long lengths of stay (Rosenstein, et al. 1990) and multiple inpatient visits (Leginski et al, 1990) have remained areas of concern. The availability of increased community resources has not always resulted in less inpatient care (Fisher, et al., 1992). Different studies have determined rates of readmission over different periods of time. Fisher, et. al (1992) found a 50% rate of readmission within 4 years of discharge. Hafemeister and Banks (1996) showed variation in readmission rates among hospitals in the same system and across years for the same hospital (p. 196). They argue that recidivism trends can be used as a performance indicator when comparing similar programs over time but warn against the misapplication to performance measurement when programs are different and when there are not several time points (p. 199).

While much research has been conducted regarding the transfer of patient care to the community, few studies have been able to specify why some areas have higher rates of readmission than others. The reduction of inpatient care in concordance with successful community treatment is a continuing challenge and needs to be informed by data-based solutions.

3. Data

The main source of data is the master demographic file for the Patient/Resident Automated Information System (PRAIS) provided by DMHMRSAS. The file contains 134,236 records created between 1978 and 1992. Each details an episode for an individual in one of Virginia's eight public adult psychiatric hospitals. Each record includes a unique patient identifier, patient demographic characteristics, administrative information collected at the patient's admission, and discharge (including beginning and ending dates for the episode), and psychiatric diagnosis codes.

Several variables were recoded into a more usable form. These variables include race, marital status, and legal status. The DSM III/IV diagnostic codes were also recoded into 13 diagnostic categories in a 2-step process: a) they were grouped into 32 diagnostic codes based on a coding scheme used by NIMH with their Client Sample Surveys; b) then they were aggregated into 13 groups based on clinical similarities and cell sizes in each group.

Histories were constructed for each individual in the data set. The beginning and ending dates of a hospital stay are given in the data. A constructed community tenure is the length of time between hospital stays or, corresponding to the last observed hospital stay, the length of time after that hospital stay until the data truncation point. Since the data seemed to be of poor quality prior to 1980,[1] we limited ourselves to the 59,497 individuals who experienced 109,333 hospital stays and 107,171 community stays after January, 1980.

Several steps were taken to minimize the effect of missing variables. If we observed a county code at either admission or discharge but not both, we set the missing

[1] Subsequently, we limited our analysis to observations after 1981. See the discussion below.

county code equal to the observed county code. Also, we assumed that the race and sex of individuals did not change over episodes and that the age of individuals changed in the expected increments over episodes. This allowed us to fill in missing race, sex, and age variables for some individuals who had more than one episode.

Next, we rejected all observations for which important explanatory variables (race, marital status, age, county code, legal status, diagnosis, or committed days) were missing. Further, we removed observations where the spell length was not positive. We also eliminated cases where the first hospital stay began before 1982 because we had concerns regarding the completeness of data obtained during the early years of a newly automated data system and differences in the earlier diagnosis coding scheme. As of 1981, all hospitals were instructed to use ICD9 diagnoses in their automated systems. When the automated system was updated in the early 1990s, the DSM III/ DSMIV became the diagnostic classification system. The DSM III/IV system was developed to be congruent with the ICD9 system. We made minor translations from DSM to ICD9 prior to aggregating into the 13 large diagnostic grouping categories. A detailed account of the impact of these selection criteria on the size of the data set is provided in Table 1.

After these rejections, the number of hospital stays fell to 61,648, and the number of constructed community stays fell to 70,051.[2,3] These stays represent information on

[2] Note that missing diagnosis codes were responsible for eliminating 19,292 hospital stays and (only) 5,467 community stays. We assumed that, if the diagnosis code was missing for the hospital stay prior to a community stay, then we would first look to the hospital stay subsequent to the community stay before rejecting the community stay. This is the primary reason that the number of community stays exceeds the hospital stays.

[3] The two major sources of missing data were missing diagnostic codes and data prior to 1981. There is no reason for us think that rejecting data prior to 1981 would bias results if we limit our analysis to the period after 1981. Missing diagnostic codes are unlikely to have a strong selection effect. Besides, there is no obvious way to correct for them especially given the poor correlation between observed diagnostic codes and other exogenous variables.

58,821 people with an average 1.05 hospital stays and 1.19 community stays. In most cases, our standards resulted in some episodes being dropped for each individual, but not all of his or her episodes. In fact, 90,869 episodes were discarded without our having to lose any people. These observations make up the data set used in this study. Table 2 indicates the portion of episodes that were right censored. Not surprisingly, a large proportion of the community stays are right censored. A much smaller number of the hospital stays are censored. The mean and median observed hospital lengths of stay were 95.4 days and 27 days, respectively, and the mean and median observed community tenures were 1,348.4 days and 637 days respectively.

As a final step, the data was supplemented by information from the Area Resources File (ARF), provided by the Bureau of Health Professions. This file contains county and city aggregate data on variables such as patient care psychiatrists, nurses, median education, nursing home and hospital beds, prison population,[4] percent urban, and percent black. Most of these variables were normalized by county population. Most variables in the ARF are reported only on a decennial basis at 1980 and 1990. Interpolation is used to obtain values away from those dates. We had concerns that some 'zeros' in the data were really either missing variables or the result of sampling problems in rural counties. Nevertheless, there was no better alternative data source with similar information.

Table 3 provides names and definitions for the variables. Table 4 presents the means and standard deviations for the explanatory variables from the ARF file and the PRAIS file for both hospital stays and community stays. There is little difference

[4] Prison population figures for 1980 actually were obtained from the 1980 Census STF4A file.

between the means of the explanatory variables for hospital stays and community stays. The 'typical' individual in the data is a white, unmarried male in his late thirties, although there are a disproportionately large number of blacks in the data relative to the population as a whole. The most commonly assigned diagnosis is schizophrenia followed by bipolar disorders and then other depressive disorders.

Several Kaplan-Meier survival curves were generated to evaluate the average effect of various exogenous characteristics on the length of hospital stays and community stays. Figure 1 demonstrates the average effect of race on community tenure and hospital stays. There is no difference in community tenure based on race until approximately 125 days in the community. At that time, individuals who are black have a slightly lower probability of remaining in the community. In contrast, hospital stays differ in length by race. At all points in time, people who are white leave the hospital faster. People who are black have about a 20% chance of remaining in the hospital at 100 days compared to 18% for those who are white. After 400 days, there is little difference between races in the probability of remaining in the hospital.

Figure 2 contrasts the survivor functions for hospital stays for individuals from rural versus urban areas. Individuals from rural areas leave the hospital faster than those from urban areas. For example, there is about a 30% probability that someone from a rural area will remain in the hospital at 50 days and a 35% probability that someone from an urban area will remain in the hospital. In contrast, those in rural areas are more likely to remain in the community longer. At 50 days, approximately 90% of those from rural areas will remain in the community compared to 85% of those from more urban areas.

Figure 3 provides the Kaplan-Meier survivor curves for hospitals stays, disaggregated by a selected set of diagnoses at admission. The curves indicate that, on average, individuals with dementia stay in the hospital the longest (66% chance that the person is in the hospital for more than 100 days), followed by schizophrenia (33%) and` depression (15%), and finally individuals being treated for substance abuse (11%) are in the hospital for the shortest period of time. The analogous survivor curves for community stays, disaggregated by discharge diagnosis, indicate that individuals with dementia stay in the community the longest time (96% chance that the person is in the community for more than 100 days), followed by depression (88%) and substance abuse (88%), and finally individuals suffering from schizophrenia (86%) are in the community for the shortest period of time. The ordering of these curves, especially the hospital stay survivor curves, is consistent with the relative severity of the disorders.

4. Econometric Specification

Let t_{ij}^h be the length of the *j*th psychiatric hospital stay for individual *i*, and let t_{ij}^c be the length of the *j*th community stay for individual *i*. Assume that *i* has n_i^h hospital stays and n_i^c community stays. Let $d_{ij}^h = 1$ if the *j*th hospital stay was censored, and $d_{ij}^h = 0$ otherwise. Define d_{ij}^c analogously for community stays. Let X_{ij}^h be a set of explanatory variables for the *j*th hospital stay, and let X_{ij}^c be a set of explanatory variables for the *j*th community stay. The conditional hazard rate at time τ is modeled as

$$\lambda^k\left(\tau \middle| \varepsilon_{ij}^k\right) = \exp\left\{X_{ij}^k \beta^k + g(\tau) + \varepsilon_{ij}^k\right\} \tag{1}$$

for $k = h, c$. This is the standard proportional hazard model (Cox 1972). The baseline hazard $g^k(\tau)$ is modeled as a piecewise linear spline function (Meyer 1990):

$$g^k(\tau) = \sum_{l=0}^{m-1} \underline{g}_l^k \left[\underline{\tau}_{l+1} - \underline{\tau}_l \right] + \underline{g}_m^k \left[\tau - \underline{\tau}_m \right] \tag{2}$$

for $\underline{\tau}_m \le \tau \le \underline{\tau}_{m+1}$. The $\underline{\tau}_l$ variables are nodes, and each $\underline{g}_l^k$ variable is a slope between $\underline{\tau}_{l+1}$ and $\underline{\tau}_l$. We fix the nodes and estimate the slopes. The last term in equation (1), ε_{ij}^k, measures unobserved heterogeneity due to covariates not observed in the data. The econometrics literature always assumes the existence of unobserved heterogeneity (e.g., Heckman and Singer 1984). Outside econometrics, demographers and statisticians such as Vaupel, Manton, and Stallard (1979), Hougaard (1984, 1986a, 1986b, 1987), and Costigan and Klein (1993) discuss unobserved heterogeneity and call it "frailty." Gordon (1996) allows the unobserved heterogeneity to interact with the covariates in a cure mixture model of death from breast cancer. An implication of the identification proof in Elbers and Ridder (1982) is that Gordon's model is identified only because it assumes a flat baseline hazard function.

In this paper, the unobserved heterogeneity ε_{ij}^k takes on different forms. In the simplest case, we assume there is no unobserved heterogeneity. In the next case, we assume ε_{ij}^k does not vary over *j* (stays of a particular type) but is independent over *k* (stay type) and *i* (individuals). In the most general case, we still assume ε_{ij}^k does not vary over *j* but allow ε_{ij}^h and ε_{ij}^c to be correlated. Consider, for example, a hypothetical person i with two hospital stays and two community stays. Then the unobserved heterogeneity terms associated with the four episodes would be $(\varepsilon_i^h, \varepsilon_i^c, \varepsilon_i^h, \varepsilon_i^c)$' with $\begin{pmatrix} \varepsilon_i^h \\ \varepsilon_i^c \end{pmatrix} \sim N\left[0, \sigma^2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]$. In all cases where there is unobserved heterogeneity, we use a

4-point Gaussian quadrature approximation to the normal distribution following Lillard (1993).[5] We also tried using a generalization of the 2-point discrete distribution approximation described in Heckman and Singer (1984) allowing for unspecified correlation but found convergence properties of the optimization algorithm were not as nice.

The log likelihood of observing a set of episode lengths, $\left\{t_{ij}^{h}\right\}_{j=1}^{n_i^h}$ and $\left\{t_{ij}^{c}\right\}_{j=1}^{n_i^c}$, is approximated as

$$L_i = \sum_{e^h}\sum_{e^c}\prod_{k=h,c}\prod_{j=1}^{n_i^h}\left[\lambda^k\left(t_{ij}^k \middle| e^k\right)\right]^{1-d_{ij}^k} \exp\left\{-\int_0^{t_{ij}^k}\lambda^k\left(s \middle| e^k\right)ds\right\}\Pr\left[\varepsilon_{ij}^h = e^h, \varepsilon_{ij}^c = e^c\right] \qquad (3)$$

where e^k takes on values implied by the covariance matrix of $\left(\varepsilon^h, \varepsilon^c\right)$ and the 4-point Gaussian quadrature approximation. The log likelihood function is

$$L = \sum_{i=1}^{N}\log L_i \ . \qquad (4)$$

It is maximized over $\theta = \left(\beta^h, \beta^c, \underline{g}^h, \underline{g}^c, \Omega\right)$ where Ω is the covariance matrix of $\left(\varepsilon^h, \varepsilon^c\right)$.

For most of the analysis, the set of explanatory variables used will consist of those listed in Table 3. In addition, there are two spike variables. The first is equal to 1 for the 3 days before and the 2 days after one's committed days for hospital stays. This allows for the high hazard rate when committed days run out. The other is equal to 1 for the first 7 days of the hospital stay for those admitted under a temporary detention order (TDO).

[5] Lillard (1993) and Sickles and Taubman (1986) find that the improvement in precision associated with increasing the number of points in the quadrature procedure past four is negligible.

The nodes for the baseline are 7 days, 15 days, 31 days, 50 days, and 60 days. Thus we estimate six $\underline{g}^k$ slopes for each type of episode.

Some of the analysis is presented in terms of the probability of remaining (either in the hospital or the community) at time t. This probability, called the survivor function, is defined in terms of the hazard rate as

$$S_{ij}^k(t) = \int \exp\left\{-\int_0^t \lambda^k\left(s\middle|\varepsilon_{ij}^k\right)ds\right\}\phi\left(\varepsilon_{ij}^k\right)d\varepsilon_{ij}^k \tag{5}$$

where $\phi\left(\varepsilon_{ij}^k\right)$ is the marginal density of ε_{ij}^k.

It is well known that discarding the first left-censored episode leads to inconsistent parameter estimates when there is unobserved heterogeneity (Heckman and Singer 1985 and Lancaster 1990). This occurs because one is undersampling long episodes relative to their distribution in the population. We choose to ignore this problem for three reasons. First, we have a very long panel of data, so inconsistency due to left censoring is probably minimal. Second, the first (left-censored) community stay corresponds to an individual before he or she became severely mentally ill and thus corresponds to a different population than the one we are interested in. Third, attempts to control for left censoring in a shorter panel in Holt, Merwin, and Stern (2001) suggest that left censoring has little effect on other parameter estimates for this sample.

5. Results

Parameter estimates are reported in Tables 5, 6, and 7. Table 5 provides parameter estimates when no unobserved heterogeneity is modeled, Table 6 when unobserved heterogeneity is correlated across episodes of the same type but is uncorrelated across episodes of different types (Ω is diagonal), and Table 7 when Ω is a general covariance matrix. With a few exceptions, the reported numbers are estimates of the derivative of the log hazard with respect to the relevant variable. For example, in Table 5, the coefficient on FEMALE is -0.104 for hospital stays. This implies that, holding other characteristics constant, men leave the hospital 10.4% faster than women. The exceptions are the $\underline{g}^k$ (SLOPE) coefficients (defined in equation (2)) and the σ (SIGMA) and ρ (CORR) coefficients; SIGMA is the standard deviation of the associated unobserved heterogeneity error (the square root of the relevant diagonal term in Ω), and CORR is the correlation.

One can see some common trends by comparing the three tables. As expected there is a negative duration dependence bias when unobserved heterogeneity is ignored (the SLOPE coefficients become more positive when going from Table 7 to Table 6).

Almost none of the other parameter estimates change sign across the three tables (BLACK is the exception), but, in many cases, magnitudes change across the tables. For example, many of the calendar effects and the NHSPS change magnitude significantly. Comparison of the log likelihoods implies that allowing for uncorrelated unobserved heterogeneity is statistically significant for both hospital episodes and community episodes. Allowing for correlated unobserved heterogeneity is also statistically significant.[6] We spend the rest of our discussion on the estimates in Table 5 because these are the best estimates.

The results show that blacks spend less time in the community {before being readmitted to the hospital}; they leave the community 8% faster than whites. There is no difference in how fast blacks get out of the hospital. Women stay in hospitals longer but leave the community at the same rate as men. There are eight hospital dummy variables. Hospital 8 (HOSP8) is significantly different from the other seven hospitals in that people stay in Hospital 8 longer and then stay in the community longer.[7] From 1984-89, clients were staying in the hospital longer than in 1982-83 but were also staying in the community longer from 1986-89. This trend reverses in 1990-92 with faster discharges from hospitals but with quicker readmissions. This could be an artifact regarding the ending of the data in 1992. Alternatively, increased community funding in the late eighties may have influenced the trend in 1990-92, thereby allowing hospitals to discharge patients to community placements more quickly. However, additional community resources apparently were not enough to prevent rapid readmission. The calendar year effects suggest a significant increase in hazard rates into and out of the hospital in 1990. It is surprising that there was not more of a steady increase in the hospital hazards and decrease in the community hazards in light of the consistent pressure to reduce the use of state inpatient care and instead substitute community care. This trend suggests the community resources were not intense enough to substitute for inpatient care.

[6] The log likelihood for the combined model with correlation is -5.582 E+5, which is significantly more than the sum of log likelihoods for the hospital stay and community stay models when there is no correlation: -2.980E+5 - 2.920E+5 = -5.900E+5.

Figure 4 shows that there has been a decline in the total number of admissions to state psychiatric hospitals over the relevant period.[8] However, our estimates of year effects suggest no such decline. In particular, there is no significant trend in the year dummies.[9] There are two possible reasons for this. First, since we condition on having a first admission, any changes caused by changes in rates of first admissions would not be captured in our estimates. Figure 4 shows that part of the decline in total admissions is due to a decline in first admissions. Second, there could be a change in the composition of first admission patients over time. Figure 5 shows that patients with schizophrenia declined as a proportion of total patients, and our results show that schizophrenic patients have longer than average stays in the hospital. Thus, both effects due to first admissions help to explain long-run trends in hospital stays, while behavior once in the hospital has little effect.

Married people are discharged faster from hospitals and stay in the community longer. Older people stay in the hospital longer, but also remain longer in the community. Employed people leave the hospital faster and remain in the community longer. The diagnosis estimates suggest that people with alcohol, substance abuse, and adjustment disorders have high hazard (or exit) rates out of the hospital, while people with schizophrenia, schizoaffective disorder, paranoia, other psychotic disorders, and bipolar disorders have low hazard rates out of the hospital relative to dementia. People with dementia have unusually low hazard rates out of the hospital. These results are consistent with our notions of illness severity.

[7] We have no way of identifying the eight hospitals from each other.
[8] Figure 4 uses all of the data including those observations with missing values that were not used in the estimation procedure.

We have two measures of previous hospital history. The longer the length of the previous hospital stay (LNPSP), the longer the hospital stay. An illness severity argument (longer hospital stays are associated with sicker people) is a possible explanation. This might help to explain why the severity of illness effect in the unobserved heterogeneity is dominated by other effects as discussed below. An increase in the number of previous hospital stays (NHSPS) also causes longer hospital stays and shorter community stays. This variable seems to be capturing some dimension of severity although a different dimension than in the unobserved heterogeneity. In particular, a large value of ε_c and a small value of ε_h (person with severe illness) would not have a large value of NHSPS because hospital stays would be long; one needs both ε_c and ε_h to be large in order for NHSPS to be large. Involuntarily admitted clients leave the hospital faster but also leave the community faster. Those admitted from jail also leave the hospital faster but are not readmitted faster.

The community characteristics have mixed effects, and no effects are very powerful. Clients from areas with a higher number of hospitals and nursing home beds leave the hospital more slowly, and those with high availability of nursing home beds remain in the community longer. People from more urban areas stay in the hospital longer, but there is no difference in readmission. There are uneven discharge and readmission rates over time.

The slope estimates suggest moderate deviations from a constant baseline hazard. The most prominent effects are the slowdown of the hazard rate at the beginning of community stays and the speedup of the hazard rate at the beginning of hospital stays.

---

[9] Had the year dummies been decreasing over time, that would have explained declining admissions.

We believe this is related to the absence of variables measuring care and lifestyle information about the stay in the community. For example, information on living arrangements, relationships, finances, and health and mental health utilization would be desirable.

There are two SIGMA estimates representing the standard deviation of the unobserved heterogeneity components. Both are large, implying that unobserved characteristics such as severity of illness and calendar effects are important determinants of hospital and community hazard rates. We estimated a correlation of 0.253, which is consistent with severity's being a large component of heterogeneity. However, when we estimated the model with correlated heterogeneity terms, the parameter estimates did not change significantly. In previous analyses on subsets of the data over shorter periods of time (Holt, Merwin, Stern, 2001), we estimated negative correlation and interpreted that to mean severity of illness was a major component of unobserved heterogeneity. Increased severity causes one to remain longer in the hospital and shorter in the community. Our positive estimate with a longer panel suggests there is a significant time component in unobserved heterogeneity, not captured by the calendar time dummies, that dominates the severity effect. In particular, there are some patients who have one relatively long hospital stay followed by a long community stay, while there are others with frequent short hospital stays. Our results suggest that the variance of frequency in and out is changing over time. This can be captured by calendar time dummies only to the extent it is reflected in mean hazard rates.

6. Conclusion

Our results show that information on client characteristics available from inpatient stay records is indeed useful in predicting not only the length of inpatient stay but also the length of the subsequent community stay after hospitalization. Our findings demonstrate that client characteristics may be used to target increased discharge planning for those at greater risk of rapid readmission to inpatient care. Correlation across observed and unobserved factors affecting length of stay has significant effects on the measurement of relationships between individual factors and lengths of stay. Thus, it is important to control for both observed and unobserved factors in estimation.

Patient characteristics influence the time spent in the hospital as well as the time spent in the community. Of note is the fact that African American clients have a shorter stay in the community before readmission than do white clients. The reasons for this should be explored in future studies and should include the availability of needed community care. People from rural areas are in the hospital for shorter periods of time but are readmitted at the same rate as people from urban areas. Finally, when patient and hospital characteristics were also considered, community characteristics provided little information about the length of hospital stay or community tenure. Implications for further research include enhancing the data set with more detailed information about patients' lives in the community. In particular, data regarding the use of {outpatient} mental health services following discharge from the hospital could reveal how such services affect rates of readmission.

Over the decade of the 1980s, Virginia observed a decline in admissions to state psychiatric hospitals. Our results suggest this is due to a decline in the rate of first admissions and in the composition of patients admitted for the first time. Further, our analysis suggests that there was little change in behavior in and out of hospitals conditional on having a first admission.

References

Costigan, T. and J. Klein (1993). "Multivariate Survival Analysis Based on Frailty Models." (ed.) A. Basu, Advances in Reliability. New York: North-Holland, pp. 43-58.

Cox, D. R. (1972). "Regression Models and Life Tables." Journal of the Royal Statistical Society, B. 34: 187-220.

Elbers, C. and G. Ridder (1982). "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model." Review of Economic Studies. 49(3): 403-410.

Fisher, W.H., J.L. Geller, F. Altaffer, and M.B. Bennett (1992). "The Relationship Between Community Resources and State Hospital Recidivism." American Journal of Psychiatry. 149(3): 385-390.

Gordon, N. (1996). "Cure Mixture Models in Breast Cancer Survival Studies." In (ed.) N. Jewell, A. Kimber, M. Lee, and G. Whitmore, Lifetime Data: Models in Reliability and Survival Analysis. Dordrecht: Kluwer Academic Publishers, pp. 107-112.

Hafemeister, T.L., and S. M. Banks (1996). "Methodological Advances in the Use of Recidivism Rates to Assess Mental Health Treatment Programs." The Journal of Mental Health Administration, 23(2):190-206.

Heckman, J. and B. Singer (1984a). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." Econometrica. 52(2): 271-320.

Heckman, J. and B. Singer (1985). "Social Science Duration Analysis." in Longitudinal Analysis of Labor Market Data. (eds.) J. Heckman and B. Singer. Cambridge: Cambridge University Press.

Holt, F., Merwin, E., Stern, S. (2001) "The Lengths of Psychiatric Hospital Stays and Community Stays." Virginia Economic Journal. Forthcoming.

Hougaard, P. (1984). "Life Table Methods for Heterogeneous Populations: Distributions Describing the Heterogeneity." Biometrika. 71(1): 75-83.

Hougaard, P. (1986a). "A Class of Multivariate Failure Time Distributions." Biometrika. 73(3): 671-678.

Hougaard, P. (1986b). "Survival Models for Heterogeneous Populations Derived from Stable Distributions." Biometrika. 73: 387-396.

Hougaard, P. (1987). "Modelling Multivariate Survival." Scandanavian Statistical Journal. 14: 291-304.

Lancaster, T. (1990). The Econometric Analysis of Transition Data. Cambridge: Cambridge University Press.

Leginski, W.A., R.W. Manderscheid, and P.R. Henderson (1990). "Patients Served in State Mental Hospitals: Results from a Longitudinal Data Base," in NIMH, Mental Health, United States 1990. R.W. Manderscheid, and M.A. Sonnenschein, eds., DHHS Pub. No. (ADM)90-1708. Washington, D.C.: Supt. of Documents, U.S. Government Printing Office.

Lillard, L. (1993). "Simultaneous Equations for Hazards: Marriage Duration and Fertility Timing." Journal of Econometrics. 56: 189-217.

Meyer, B. (1990). "Unemployment Insurance and Unemployment Spells." Econometrica. 58(4): 757-782.

Rosenstein, M.J., L.J. Milazzo-Sayre, and R.W. Manderscheid (1990). "Characteristics of Persons Using Specialty Inpatient, Outpatient, and Partial Care Programs in 1986," in NIMH, Mental Health, United States 1990. R.W. Manderscheid, and M.A. Sonnenschein, eds., DHHS Pub. No. (ADM)90-1708. Washington, D.C.: Supt. of Documents, U.S. Government Printing Office.

Sickles, R. and P. Taubman (1986). "An Analysis fo the Health and Retirement Status of the Elderly." Econometrica. 54(6): 1339-1356.

Vaupel, J., K. Manton, and E. Stallard (1979). "The Impact of Heterogeneity in Individual Frailty on the Dynamics of Mortality." Demography. 16: 439-454.

Witkin, M.J. J.E. Atay, A.S. Fell, and R.W. Manderscheid (1990). "Specialty Mental Health System Characteristics," in NIMH, Mental Health, United States 1990. R.W. Manderscheid, and M.A. Sonnenschein, eds., DHHS Pub. No. (ADM)90-1708. Washington, D.C.: Supt. of Documents, U.S. Government Printing Office.

Table 1

Missing Variable Analysis

Cause	Hospital Stays		Community Stays	
	# Obs Lost	Cumulative	#Obs Lost	Cumulative
Race	963	963	944	944
Marital Status	1646	2609	1593	2537
Age	92	2701	51	2588
County Code	233	2934	22	2610
Legal Status	28	2962	30	2640
Diagnosis	19292	22254	5467	8107
Miscellaneous	26	22,280	422	8529
Data from 80-81	25,405	47,685	28,591	31,120

Table 2

Dependent Variable Sample Sizes

	Hospital Stays	Community Stays
Without Censoring	60047	38426
Right Censored	1601	31625
Total	61648	70051
Mean Episode Length (days)	95.4	1348.4

Table 3

Variable Name Definitions

BLACK	Respondent is black
FEMALE	Respondent is female
HOSPn	Hospital dummies (n=1,2,..8)
Y8283	Episode began in 1982-1983
Y8485	Episode began in 1984-1985
Y8687	Episode began in 1986-1987
Y8889	Episode began in 1988-1989
Y9092	Episode began in 1990-1992
MARRY	Respondent is married
AGE	Respondent's age
DEMENTIA	Diagnosis is dementia
SUBSTABU	Diagnosis is substance abuse
ALCOHOL	Diagnosis is alcohol abuse
ORGANIC	Diagnosis is organic disorder
SCHIZ	Diagnosis is schizophrenia
SCHIZAFF	Diagnosis is schizo-affective disorder
PARANOID	Diagnosis is paranoid disorder
OTHPSYTIC	Diagnosis is other psychotic disorder
BIPOLAR	Diagnosis is bipolar disorder
DEPRESS	Diagnosis is depression
PERSNLTY	Diagnosis is personality
ADJUST	Diagnosis is adjustment disorder
OTHERDIAG	Diagnosis is other
EMPLOY	Respondent is employed
EMPMIS	Respondent's employment status is missing
COMMDYS	Number of committed days
LNPSP	Respondent's longest previous hospital stay
NHPSP	Number of previous hospital stays for the respondent
LGST	Respondent entered the hospital involuntarily
JAIL	Respondent came from jail
CTY-MD	Per-capita medical doctors in county/city
CTY-RN	Per-capital RN's in county/city
CTY-LPN	Per-capita LPN's in county/city
CY-HOSP	Per-capita hospitals in county/city
CTY-PCI	Per-capita income in county/city
CTY-URB	Percent of county/city that is urban
CTY-BLK	Percent of county/city that is black
CTY-JAIL	Percent of county/city in jail
CTY-NHB	Per-capita nursing home beds

Table 4

Explanatory Variables

	Hospital Stays		Community Stays	
	Mean	Std Dev	Mean	Std Dev
BLACK	.32	.47	.33	.47
FEMALE	.39	.49	.37	.48
Y8283	.16	.37	.20	.4
Y8485	.19	.39	.19	.39
Y8687	.19	.39	.18	.38
Y8889	.19	.39	.19	.39
Y9092	.28	.44	.25	.43
MARRY	.15	.36	.16	.37
AGE	38.56	14.90	38.74	15.35
DEMENTIA	.02		.02	
SUBSTABU	.04	.19	.04	.20
ALCOHOL	.15	.36	.15	.36
ORGANIC	.06	.24	.07	.26
SCHIZ	.24	.42	.25	.43
SCHIZAFF	.08	.27	.09	.29
PARANOID	.01	.09	.01	.10
OTH-PSYTIC	.05	.22	.04	.18
BIPOLAR	.10	.29	.10	.30
DEPRESS	.10	.30	.08	.27
PERSNLTY	.03	.17	.04	.19
ADJUST	.06	.23	.07	.26
OTHERDIAG	.06	.25	.04	.20
COMMDYS	853.31	351.60		
LNPSP			198.87	1080.58
NHPSP	3.38	3.89	3.29	3.80
LGST	.70	.46	.60	.49
CTY-MD	.15	.14	.15	.13
CTY-RN	.61	.24	.60	.24
CTY-LPN	.19	.10	.20	.11
CTY-HOSP	.002	.002	.002	.002
CTY-PCI	15.11	5.45	14.77	5.36
CTY-URB	.56	.39	.56	.37
CTY-BLK	.20	.15	.20	.16
CTY-JAIL	.53	1.15	.55	1.15
CTY-NHB	.46	.41	.45	.37

Table 5

Parameter Estimates with Correlated Unobserved Heterogeneity

		Community Stay				Hospital Stay		
Name		Value	Std.Error	t-statistic		Value	Std.Error	t-statistic
BLACK		0.076	0.018	4.22		-0.019	0.017	-1.11
FEMALE		0.012	0.017	0.75		-0.188	0.015	-12.15
HOSP1		-4.507	0.068	-66.46		-4.603	0.061	-75.76
HOSP2		-4.294	0.074	-58.44		-4.480	0.065	-68.53
HOSP3		-4.271	0.070	-60.81		-4.842	0.062	-77.86
HOSP4		-4.458	0.069	-64.41		-4.883	0.061	-79.93
HOSP5		-4.495	0.070	-64.27		-5.067	0.060	-83.95
HOSP6		-4.472	0.070	-64.02		-4.948	0.064	-77.61
HOSP7		-4.365	0.074	-59.05		-4.945	0.065	-75.97
HOSP8		-5.079	0.114	-44.41		-5.683	0.082	-69.66
Y8485		-0.017	0.019	-0.90		-0.062	0.022	-2.89
Y8687		-0.082	0.021	-3.87		-0.114	0.023	-4.99
Y8889		-0.103	0.025	-4.20		-0.121	0.025	-4.78
Y9092		0.206	0.030	6.89		0.124	0.029	4.22
MARRY		-0.140	0.021	-6.78		0.373	0.019	19.55
AGE		-0.011	0.000	-20.35		-0.019	0.000	-41.95
SUBSTABU						1.613	0.049	32.77
ALCOHOL						2.200	0.044	50.21
ORGANIC						0.895	0.044	20.51
SCHIZ						0.635	0.043	14.80
SCHIZAFF						0.741	0.048	15.48
PARANOID						0.799	0.070	11.42
OTHPSYTIC						0.823	0.047	17.43
BIPOLAR						0.954	0.046	20.68
DEPRESS						1.175	0.045	26.22
PERSNLTY						1.159	0.052	22.32
ADJUST						1.596	0.046	34.46
OTHERDIAG						1.026	0.048	21.19
EMPLOY		-0.253	0.107	-2.37		0.758	0.055	13.86
EMPMIS		-0.217	0.039	-5.60		0.624	0.021	30.04
LNPSP		0.000	0.000	-11.81		-0.054	0.003	-16.15
NHSPS		0.105	0.003	34.35		-0.176	0.018	-10.08
LGST		0.211	0.017	12.63		0.569	0.017	34.40
CTY-MD		0.175	0.089	1.97		0.368	0.070	5.24
CTY-RN		-0.164	0.047	-3.52		0.065	0.042	1.55
CTY-LPN		0.872	0.094	9.29		0.136	0.085	1.59
CTY-HOSP		-1.814	4.596	-0.39		-11.440	3.944	-2.90
CTY-PCI		0.000	0.003	0.01		-0.007	0.002	-3.32
CTY-URB		0.011	0.030	0.36		-0.272	0.027	-10.10
CTY-BLK		-0.249	0.060	-4.14		-0.064	0.058	-1.10
CTY-JAIL		0.011	0.007	1.70		0.015	0.005	2.88
CTY-NHB		-0.148	0.024	-6.24		-0.048	0.019	-2.49
JAIL		0.002	0.003	0.79		0.060	0.002	27.04

Parameter Estimates with Correlated Unobserved Heterogeneity (cont.)

		Community Stay				Hospital Stay		
Name		Value	Std.Error	t-statistic		Value	Std.Error	t-statistic
SLOPE 1-7 days		-0.297	0.008	-37.87		0.039	0.004	10.15
SLOPE 8-15 days		0.001	0.007	1.33		0.005	0.003	1.85
SLOPE 16-31 days		-0.031	0.004	-8.35		0.040	0.002	26.60
SLOPE 32-50 days		0.005	0.004	-1.33		-0.011	0.002	-7.34
SLOPE 51-60 days		-0.045	0.005	-9.96		-0.022	0.002	-9.26
SLOPE 61+		0.001	0.000	-435.03		0.001	0.000	-120.25

Name		Value	Std.Error	t-statistic
SIGMA1		1.105	0.019	59.41
SIGMA2		0.647	0.011	57.26
CORR		0.253	0.012	21.92

Note: DEMENTIA is a base for diagnosis, and Y8283 is a base for calendar effects.
Log-Likelihood at convergence: -5.582E+05

Table 6

Parameter Estimates with Uncorrelated Unobserved Heterogeneity

		Community Stay				Hospital Stay		
Name		Value	Std.Error	t-statistic		Value	Std.Error	t-statistic
BLACK		0.103	0.016	6.29		-0.013	0.014	-0.90
FEMALE		0.014	0.015	0.92		-0.190	0.013	-14.62
HOSP1		-4.528	0.057	-79.62		-4.431	0.041	-108.17
HOSP2		-4.293	0.061	-70.47		-4.303	0.045	-95.99
HOSP3		-4.261	0.058	-72.92		-4.623	0.042	-109.98
HOSP4		-4.471	0.058	-77.54		-4.689	0.040	-117.78
HOSP5		-4.515	0.058	-77.98		-4.895	0.040	-123.51
HOSP6		-4.507	0.058	-77.80		-4.801	0.041	-116.18
HOSP7		-4.346	0.063	-69.05		-4.708	0.047	-100.88
HOSP8		-5.135	0.105	-48.85		-5.493	0.063	-86.61
Y8485		0.002	0.017	0.10		-0.050	0.018	-2.86
Y8687		-0.069	0.019	-3.63		-0.115	0.019	-6.22
Y8889		-0.108	0.022	-5.00		-0.113	0.020	-5.56
Y9092		0.109	0.026	4.15		0.121	0.023	5.22
MARRY		-0.150	0.019	-7.76		0.354	0.016	21.83
AGE		-0.013	0.001	-27.57		-0.018	0.000	-51.26
SUBSTABU						1.447	0.040	35.86
ALCOHOL						2.010	0.035	56.96
ORGANIC						0.809	0.037	22.09
SCHIZ						0.555	0.035	15.80
SCHIZAFF						0.616	0.038	16.06
PARANOID						0.709	0.062	11.38
OTHPSYTIC						0.725	0.040	18.25
BIPOLAR						0.813	0.038	21.49
DEPRESS						1.045	0.037	28.36
PERSNLTY						0.988	0.043	22.98
ADJUST						1.451	0.038	37.89
OTHERDIAG						0.908	0.041	22.34
EMPLOY		-0.201	0.100	-2.01		0.721	0.048	15.06
EMPMIS		-0.164	0.032	-5.12		0.581	0.017	33.84
LNPSP		0.000	0.000	-9.85		-0.049	0.002	-33.32
NHSPS		0.102	0.002	50.61		-0.181	0.014	-12.64
LGST		0.191	0.015	12.85		0.507	0.014	36.80
CTY-MD		0.178	0.080	2.22		0.323	0.059	5.52
CTY-RN		-0.150	0.042	-3.57		0.044	0.034	1.29
CTY-LPN		0.787	0.084	9.41		0.304	0.069	4.39
CTY-HOSP		6.045	4.138	1.46		-12.320	3.334	-3.70
CTY-PCI		0.002	0.002	-0.78		-0.007	0.002	-3.85
CTY-URB		0.018	0.027	0.66		-0.227	0.022	-10.18
CTY-BLK		-0.213	0.053	-4.00		-0.104	0.047	-2.21
CTY-JAIL		0.011	0.006	1.72		0.022	0.004	4.95
CTY-NHB		-0.159	0.022	-7.34		-0.041	0.015	-2.68
JAIL		0.002	0.002	0.68		0.056	0.002	30.59

Parameter Estimates with Uncorrelated Unobserved Heterogeneity (cont.)

		Community Stay				Hospital Stay		
Name		Value	Std.Error	t-statistic		Value	Std.Error	t-statistic
SLOPE 1-7 days		-0.279	0.007	-40.91		0.024	0.003	7.28
SLOPE 8-15 days		0.007	0.007	1.06		0.001	0.003	-0.30
SLOPE 16-31 days		-0.029	0.003	-8.86		0.036	0.001	26.42
SLOPE 32-50 days		0.003	0.003	-0.92		-0.013	0.001	-9.65
SLOPE 51-60 days		-0.040	0.004	-10.21		-0.024	0.002	-11.56
SLOPE 61+ days		0.001	0.000	-500.37		0.001	0.000	-185.86

Name		Value	Std.Error	t-statistic		Value	Std.Error	t-statistic
SIGMA		1.139	0.016	70.43		1.100	0.011	101.04

Note: DEMENTIA is a base for diagnosis, and Y8283 is a base for calendar effects.
Log-Likelihood at convergence: -2.980D+05 (community stays)
Log-Likelihood at convergence: -2.920D+05 (hospital stays)

Table 7

Parameter Estimates with No Unobserved Heterogeneity

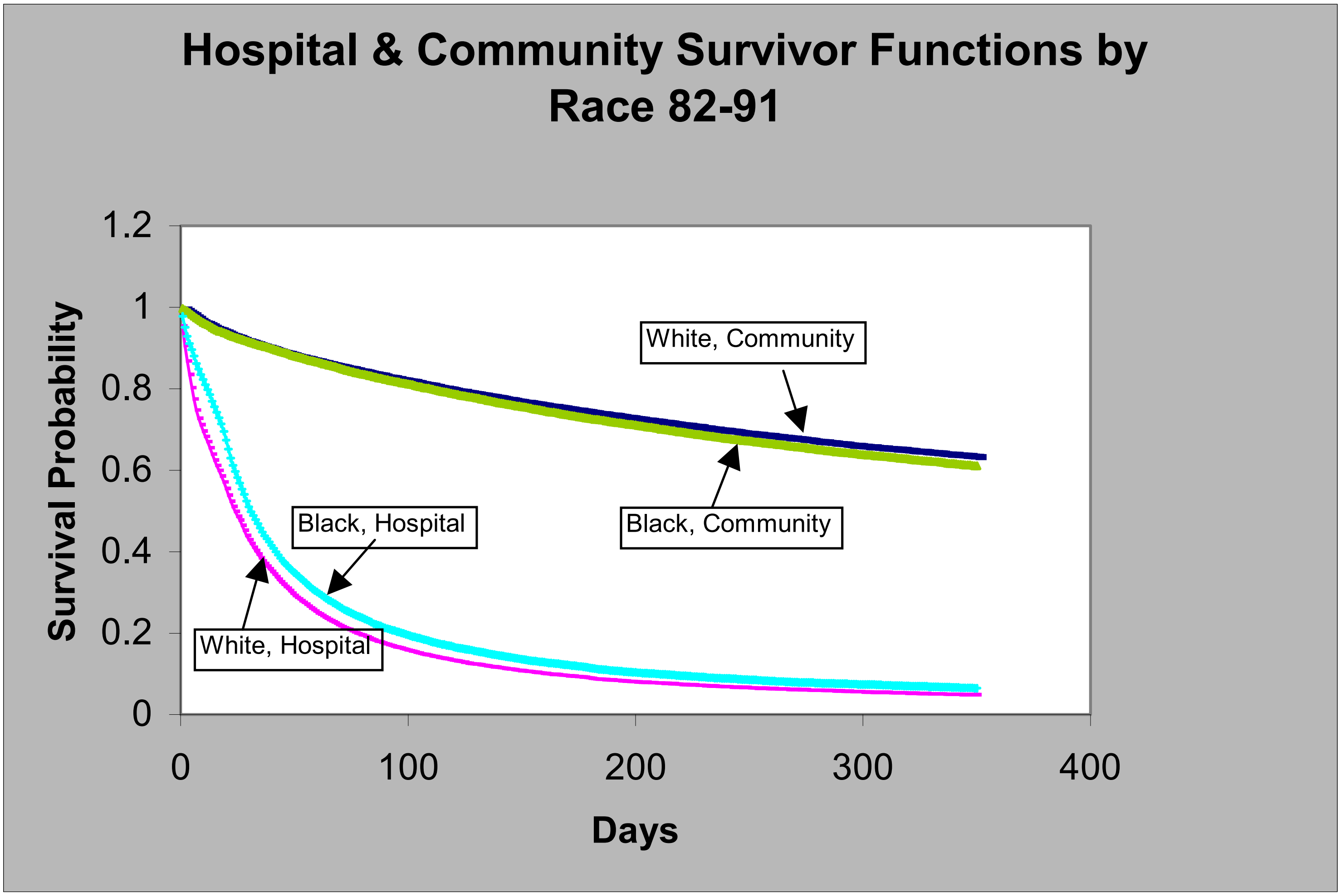
		Community Stay				Hospital Stay		
Name		Value	Std.Error	t-statistic		Value	Std.Error	t-statistic
BLACK		0.066	0.010	6.55		0.034	0.008	4.34
FEMALE		0.038	0.010	3.91		-0.104	0.007	-14.32
HOSP1		-4.162	0.046	-90.38		-4.057	0.030	-135.20
HOSP2		-4.016	0.048	-83.10		-4.011	0.033	-122.21
HOSP3		-3.913	0.047	-83.90		-4.188	0.030	-138.74
HOSP4		-4.128	0.046	-90.15		-4.277	0.029	-146.01
HOSP5		-4.156	0.047	-88.62		-4.475	0.029	-153.17
HOSP6		-4.189	0.047	-89.35		-4.387	0.031	-142.90
HOSP7		-3.990	0.050	-79.87		-4.258	0.034	-126.80
HOSP8		-4.650	0.089	-52.33		-4.857	0.047	-103.51
Y8485		0.035	0.015	2.43		-0.070	0.013	-5.35
Y8687		0.018	0.016	1.18		-0.131	0.014	-9.68
Y8889		-0.013	0.017	-0.78		-0.139	0.015	-9.50
Y9092		0.228	0.020	11.32		0.057	0.016	3.56
MARRY		-0.142	0.014	-9.92		0.290	0.010	29.14
AGE		-0.011	0.000	-37.39		-0.012	0.000	-75.04
SUBSTABU						1.134	0.031	36.20
ALCOHOL						1.687	0.028	60.57
ORGANIC						0.517	0.030	17.52
SCHIZ						0.397	0.028	14.13
SCHIZAFF						0.463	0.031	15.03
PARANOID						0.529	0.049	10.82
OTHPSYTIC						0.568	0.032	17.78
BIPOLAR						0.652	0.030	21.70
DEPRESS						0.808	0.030	27.15
PERSNLTY						0.730	0.034	21.70
ADJUST						1.114	0.030	37.28
OTHERDIAG						0.651	0.032	20.47
EMPLOY		-0.311	0.088	-3.54		0.632	0.031	20.38
EMPMIS		-0.191	0.028	-6.80		0.525	0.013	39.19
LNPSP		0.000	0.000	-14.37		-0.021	0.001	-23.95
NHSPS		0.131	0.001	169.12		-0.105	0.010	-10.23
LGST		0.204	0.012	17.06		0.346	0.010	35.36
CTY-MD		0.040	0.058	0.68		0.225	0.037	6.09
CTY-RN		-0.099	0.032	-3.07		0.035	0.022	1.61
CTY-LPN		0.664	0.062	10.70		0.076	0.043	1.75
CTY-HOSP		32.090	2.936	10.93		-6.248	2.101	-2.97
CTY-PCI		0.000	0.002	-0.04		0.004	0.001	-4.24
CTY-URB		0.031	0.020	1.58		-0.134	0.014	-9.80
CTY-BLK		-0.219	0.039	-5.61		-0.019	0.028	-0.66
CTY-JAIL		0.001	0.005	2.04		0.015	0.003	5.49
CTY-NHB		-0.113	0.017	-6.83		-0.023	0.010	-2.38
JAIL		0.004	0.002	2.22		0.041	0.001	35.35

Parameter Estimates with No Unobserved Heterogeneity (cont.)

	Community Stay			Hospital Stay		
Name	Value	Std.Error	t-statistic	Value	Std.Error	t-statistic
SLOPE 1-7 days	-0.288	0.007	-42.90	-0.026	0.003	-8.31
SLOPE 2-15 days	0.008	0.007	1.20	-0.016	0.003	-6.55
SLOPE 16-31 days	-0.032	0.003	-9.55	0.020	0.001	15.52
SLOPE 32-50 days	0.002	0.003	-0.80	-0.020	0.001	-15.13
SLOPE 51-60 days	-0.051	0.004	-13.44	-0.051	0.002	-25.96
SLOPE 61+	0.001	0.000	-581.08	0.002	0.000	-303.82

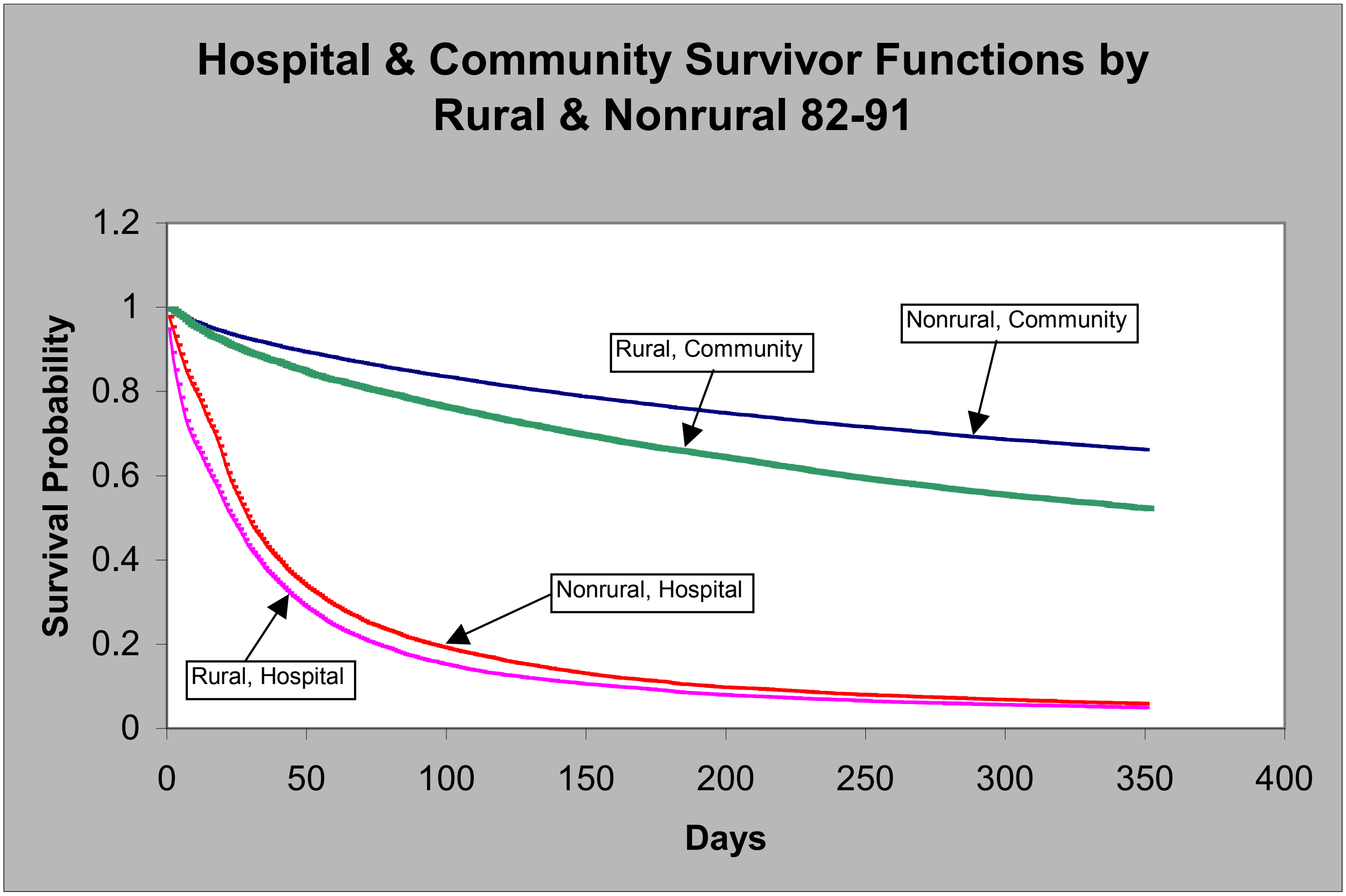
Note: DEMENTIA is a base for diagnosis, and Y8283 is a base for calendar effects.
Log-Likelihood at convergence: -2.993D+05 (community stays)
Log-Likelihood at convergence: -2.946D+05 (hospital stays)

Figure 1



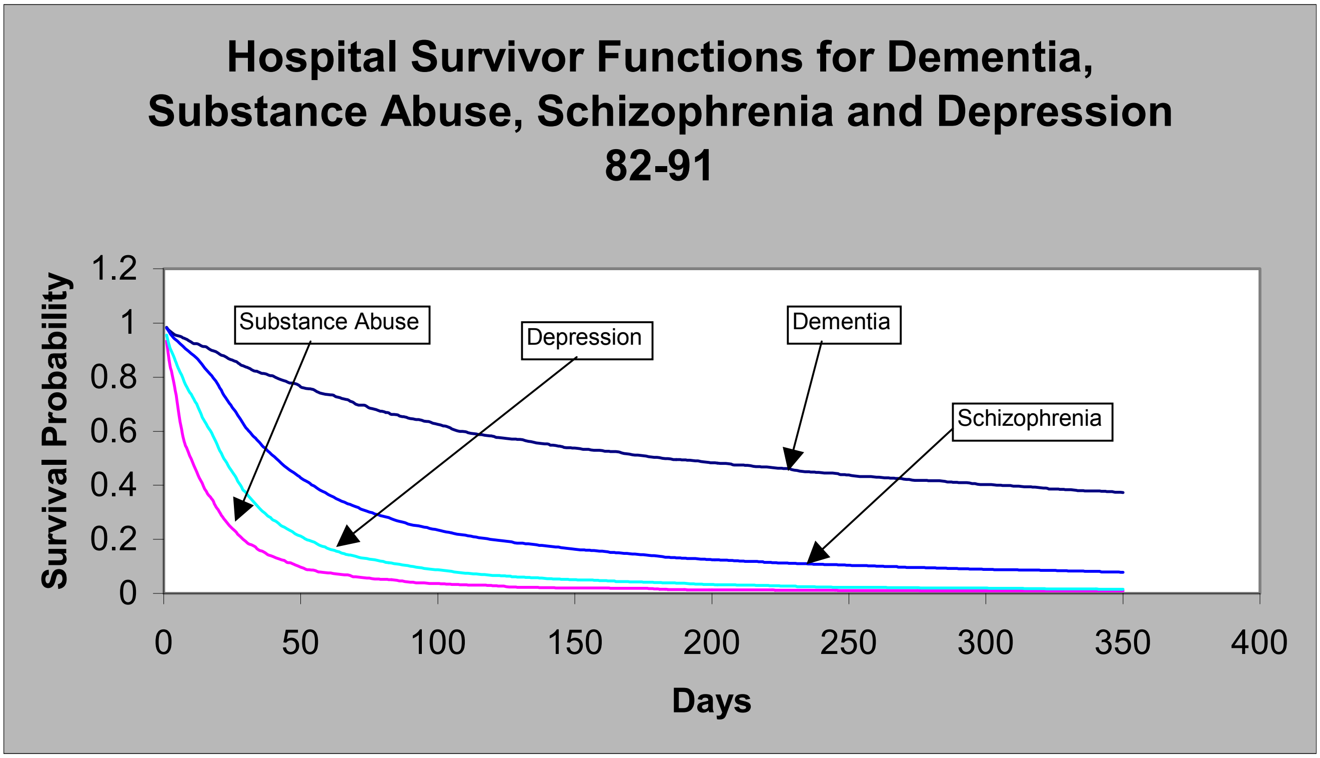
Hospital & Community Survivor Functions by Race 82-91
Survival Probability
1.2
1
0.8
0.6
0.4
0.2
0
0
100
200
300
400
Days
White, Community
Black, Community
Black, Hospital
White, Hospital

Figure 2



Hospital & Community Survivor Functions by Rural & Nonrural 82-91
Survival Probability
Days
1.2
1
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
300
350
400
Rural, Community
Nonrural, Community
Nonrural, Hospital
Rural, Hospital

Figure 3



Hospital Survivor Functions for Dementia, Substance Abuse, Schizophrenia and Depression 82-91
Survival Probability
1.2
1
0.8
0.6
0.4
0.2
0
Substance Abuse
Depression
Dementia
Schizophrenia
0
50
100
150
200
250
300
350
400
Days

Figure 4

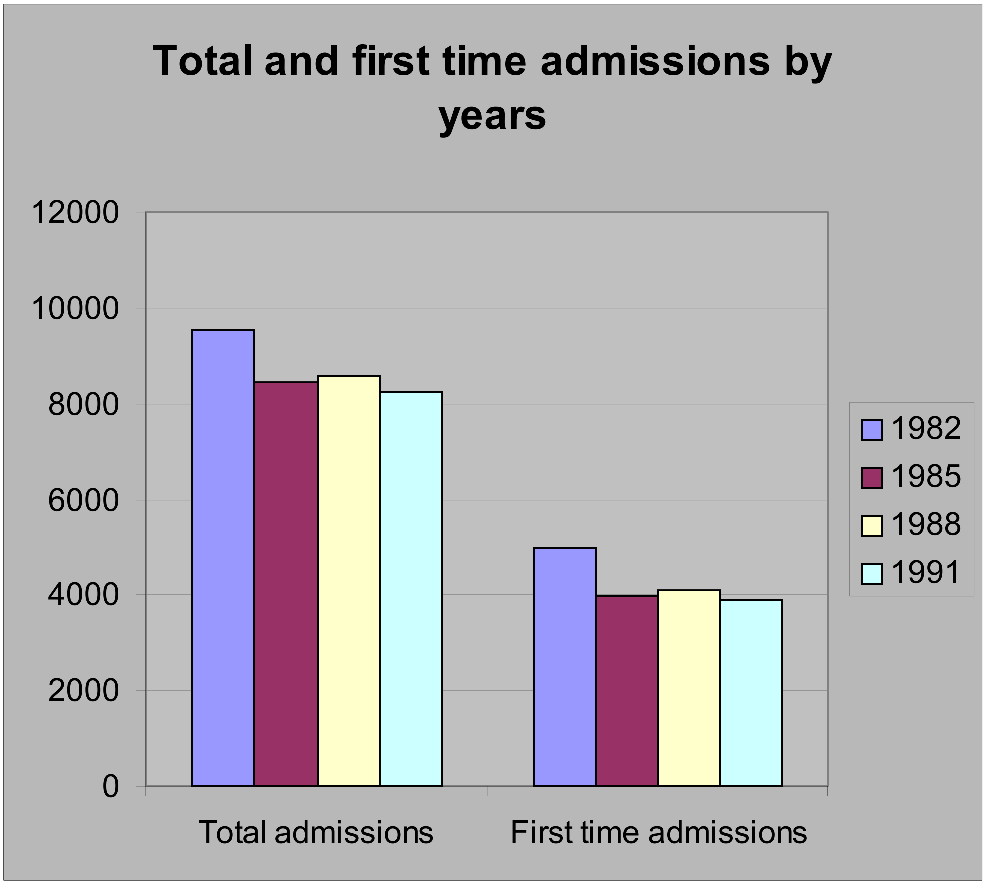
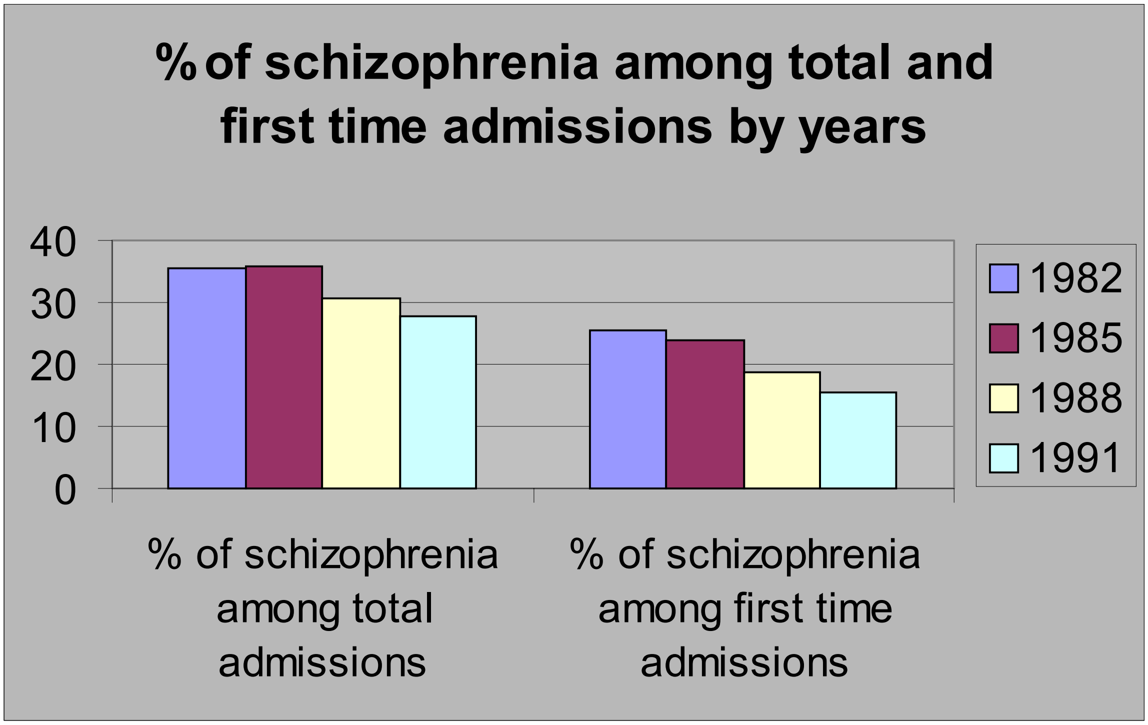

Total and first time admissions by years
12000
10000
8000
6000
4000
2000
0
Total admissions
First time admissions
1982
1985
1988
1991

Figure 5


% of schizophrenia among total and first time admissions by years
40
30
20
10
0
1982
1985
1988
1991
% of schizophrenia among total admissions
% of schizophrenia among first time admissions